A Genome-wide Screen to Identify Endoreduplication-Specific Factors

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation

with honors research distinction in Molecular Genetics in the

undergraduate colleges of The Ohio State University

by

Jonathan Lee

The Ohio State University

June 2012

Project Advisor: Professor Maki Asano, Department of Molecular Genetics

Abstract

The cell cycle process called endoreduplication is a specialized form of DNA replication. Cells do not divide and instead continually undergo G and S phase and as a result, increase their ploidy and size. Endoreduplication is important for normal development in cells such as hepatocytes, giant trophoblasts, and megakaryocytes.

However, it has also been discovered that cancer cells utilize endoreduplication in order to escape apoptosis when exposed to chemotherapeutic drugs which inhibit mitosis. Thus, it is important to identify endoreduplication specific genes in order to fully understand the regulation mechanisms which are essential for development. Moreover, our previous work and work from other labs has demonstrated that endoreduplication uses machinery as well as regulatory mechanisms which are fundamentally different from those used in mitotic replication.

In order to identify endoreduplication-specific factors, I have begun a genome-wide screening of ~12000 *Drosophila melanogaster* genes, each possessing a human homologue. The basis of the screen is a very powerful tool called the *GAL4/UAS* system, which has proven invaluable when coupled with the RNA interference (RNAi) pathway. Combining the *GAL/UAS* system with *UAS>RNAi* enables us to test the effects of downregulating each gene. The genome-wide screen consists of two concurrent screens. The first is a screen through endocycling and mitotically cycling tissue to identify endoreduplication specific factors. The second consists of screening for essential genes for the transition from mitotic replication to endoreduplication. Currently, I have screened 1280 lines and are further studying candidate genes using clonal analysis and flow cytometry techniques. I expect that one percent of the genes (~10 genes) of the 1280 will be related to endoreduplication and estimate that one percent of the *Drosophila* genome (~120) will be related to endoreduplication.

Introduction

During development, organisms use the mitotic cycle and the endocycle for different purposes. In the mitotic cycle, the cell undergoes G1, S, G2, and finally M phase to increase cell number. However, in the endocycle, the cell solely undergoes G phase then S phase and repeats this process. In other words, the cell increases its ploidy and size by skipping cell division. (1)

Our laboratory has found that the factors in endoreduplication differ from those in mitotic replication in *Drosophila melanogaster*. In mitotic replication, the Origin Recognition Complex (ORC), the funding member that initiates formation of the Pre-Replicative complex (pre-RC), is required to recruit Cell division cycle 6 (Cdc6), Double Parked (Dup), and MCM (Mini Maintenance Chromosome) to form the pre-RC. However, in endoreduplication, the Asano lab has found that the ORC is dispensable but Dup, Cdc6 and MCM are necessary. (Figure 1) (2) These results tell us that endoreduplication and mitotic replication use different machinery. As Cdc6 and Dup do not bind to the DNA on their own, the Asano lab predicts that another protein or group of proteins is responsible for initiating pre-RC formation.

Determining the initiation machinery and control mechanisms of endoreduplication is important because certain types of cancer cells are known to undergo endoreduplication when treated with anti-mitotic drugs. (3) To identify the genes involved in the control of cell type-specific endoreduplication, the factors involved in the transition from mitotic replication to endoreduplication, and genes generally involved in endoreduplication, our lab decided that a genome-wide screening was not only practical, but also essential. However, currently there is no intact organism available for a genome-wide screening in any mammalian system. In order to circumvent this problem, I am implementing *Drosophila melanogaster* as a model system and the RNAi library I am screening contains selectivity against genes all of which possess human homologues.

Findings from this project will open up an entirely new window for studying DNA replication, signaling pathways that control the transition from mitotic replication to endoreduplication during organism development, and will pave the way for another approach toward curing cancer.

Methodology

The basis for the genome-wide screening is a powerful system called the *GAL4/UAS* system coupled with the RNAi pathway. (Figure 2)

GAL4/UAS System

The *GAL4/UAS* system utilizes the transcription factor activator protein Gal4 which binds to the Upstream Activator Sequence (UAS). When the Gal4 protein binds to the UAS it activates the gene downstream of the UAS. In this sense, Gal4 protein 'drives' gene expression. (4) The *GAL4/UAS* system has several distinct advantages over other methods. The first advantage is the ability to express genes in a developmental-specific and tissue-specific manner. Transcription of Gal4 protein is under the control of a promoter which is only activated in certain subsets of fly tissues or during certain periods of development. The second benefit is that 5638 *GAL4* drivers are available in Bloomington Stock Center. Thus, one can simply choose and order the *GAL4* drivers desired for different experiments. Moreover, benefit of using the *GAL4/UAS* system is viability of transgenic flies. Ubiquitous expression of potent genes might kill a mutant fly. However, since the *GAL4* and *UAS* transgenes are separated in two different fly lines, the system is inert until when two fly lines are crossed and produce progeny which contain both transgenes. As a result, I can maintain healthy flies until I cross them to ascertain the effects of the gene of interest on the fly. Because the *GAL4/UAS* system is so versatile, it has been thoroughly

characterized and successfully tested in various *Drosophila* projects including other screenings of to identify genes involved in various cellular processes. (5) *GAL4* driver selection will be addressed below.

RNAi Pathway

The other half of the screening implements what is known as RNA interference (RNAi). RNAi is a well studied pathway originally characterized in *Caenorhabditis elegans* which moderates gene activity. (7) Activity of the RNAi pathway is controlled by the RNA-Induced Silencing Complex (RISC). (Figure 3) Trigger RNA (consisting of either dsRNA or miRNA) activates the pathway. Dicer binds and cleaves the trigger RNA into fragments called small-interfering RNA (siRNA) and the siRNA is integrated into the RISC complex. Next, using the integrated siRNA, RISC searches for and targets matching mRNA strands. Targeted mRNA by the RISC complex is translationally inhibited in two ways: either by degradation via a protein called Slicer or by continued binding of the RISC complex to the mRNA.

By combining the *GAL4/UAS* system with the RNAi pathway, I am able to downregulate gene expression in a tissue-specific manner. Using this as the basis, I plan to screen 11000 genes out of 14000 of *Drosophila*, essentially the entire genome. The transgenic lines expressing RNAi are from Vienna Drosophila RNAi Center (VDRC) and each targets one gene of the *Drosophila* genome. An UAS sequence is upstream of the RNAi which allows Gal4 to drive RNAi expression. Moreover, each gene of the 11000 targeted by RNAi has a human homologue which I am most interested in because I would like to extend the findings of my research to mammals in the future.

FACS Analysis

The FACS analysis involves dissecting ovaries from female flies, purifying their cells, and examining the DNA content profile of follicle and nurse cells. The protocol I used was developed by Dr. Brian Calvi and Dr. Mary Lilly. (14) I have adapted and conditioned certain steps of the protocol to fit my experiments. The first step is to cross the follicle cell and nurse cell RNAi candidates to their drivers (*c355>GAL4* and *VP16nos>GAL4* respectively). After 11 days, I put the flies in a cage with an apple juice-agar plate. This plate has yeast paste spread in the middle of it to stimulate oogenesis. The plate is replaced with another apple juice-agar plate spread with yeast paste twice a day until day 14. Then, my peers and I dissect out the ovaries of the progeny in Ephrussi-Beadles Ringer's (EBR) (130 m*M* NaCl, 5 m*M* KCl, 2 m*M* $CaCl_2$, 10 m*M* HEPES, pH 6.9; autoclave and store at 4° C) solution. Conditioning experiments revealed that dissecting out 40 pairs of ovaries for the follicle cells and 60 pairs of ovaries for the nurse cells yields enough cells for the FACS analysis. We place these ovaries in individual microcentrifuge tubes, suspending them in EBR. Next, we suspend the ovaries in 10 mg/ml collagenase to digest the connecting tissue of the ovarioles for 15 minutes while rotating the tubes. After digestion, we rinse the ovaries with EBR twice and resuspend the digested ovaries in nuclear isolation buffer with 1.5% NP-40 (15 m*M* Tric-HCl, pH 7.4, 60 m*M* KCl, 15 m*M* NaCl, 250 m*M* sucrose, 1 m*M* EDTA, 0.1 m*M* EGTA, 0.15 m*M* spermine, 0.5 m*M* spermidine). Then, we homogenize the ovaries, filter them twice through filter tips, and layer the samples on a sucrose step gradient which, from most dense on the bottom to least dense on the top, consists of 2.5-M, 1.6-M, and 0.8-M sucrose (dissolved in nuclear isolation buffer). Finally, we place the samples in the ultracentrifuge for 1 hour at 25000 rpm. When it is done, resuspend the pellet in nuclear isolation buffer and put DAPI in each sample.

Clonal Analysis

The clonal analysis has 3 components: a Heat Shock promoter (heat shock protein 70 aka hsp>70), FLP recombinase, the *GAL4/UAS* system. (Figure 12) hsp70 is induced to act in high-stress environments, for example under conditions where the environment is above an acceptable level for organismal development. Once induced, the gene under control of the heat shock promoter, in the case FLP recombinase and 'flip out' these sites which allows the *GAL4* driver, *Act5c>GAL4,* to be ubiquitously expressed. As the *CD2>Act5c>*Cd2 allele is linked to *UAS>RFP*, the cells in which Act5c protein is active can be distinguished by the presence of RFP. Using *UAS>RFP* allows comparison of cells expressing the RNAi to those that are not. Because the 'flip-out' event occurs in one or two cells in each salivary gland, it is possible to screen for cell autonomous interactions. In practice, the progeny of the *GAL4* driver crossed to the candidate are heat-shocked, one or two cells expressing the *UAS>RNAi* surrounded by other cells which are not can be observed. As a result, one will be able to directly compare cells expressing the candidate RNAi which will be red due to the RFP to cells which are not. It is expected that if the candidate is needed in endoreduplication, and functions cell autonomously, then the RNAi-expressing cells will be smaller in size to control cells around it. However, if the candidate is not expressed in a cell autonomous manner, the RFP minus cells surrounding the red cell will also demonstrate endoreduplication defects.

Results

Pre-screening for Gal4 Drivers

When I finished the pre-screening, I had selected three different *GAL4* drivers. They are dpp^{blk1}*>GAL4*, *c355>GAL4*, and *VP16nos>GAL4*. I chose these *GAL4* drivers because they demonstrated strict tissue specificity and because the phenotypes of the *GAL4* drivers when

expressing RNAi are easily screenable. *dpp^{blk1}>GAL4* is expressed in larval salivary glands and wing imaginal discs (in the larval wing discs which develop and become adult wings), *c355>GAL4* is expressed in ovarian follicle cells, and *VP16nos>GAL4* is expressed in germline nurse cells.

I originally aimed to distinguish factors involved specifically in endoreduplication from factors involved in both endoreduplication and mitotic replication. To do so, I needed to identify a driver which would be suitable to screen for genes which may be involved in mitosis or other processes, such as metabolism, but not involved in the endoreduplication machinery itself. I defined a 'suitable' driver as one which demonstrated tissue-specific expression and had a phenotype which one could screen quickly, spending less than ten seconds analyzing each RNAi line. The first *GAL4* drivers I screened were *GMR>GAL4* and *eyeless>GAL4*. Both are expressed in mitotically replicating eye imaginal discs and have been utilized in multiple screens by other laboratories. (8, 9) Although information regarding these drivers is available in stock center databases, I decided to test them because *UAS>RNAi* expression may potentially change the *GAL4* driver activity. To do so, I crossed them to six *UAS>RNAi* lines specific for factors known to be involved with endoreduplication and genes involved with *Drosophila* development. I used *UAS>fru*, a gene involved with male courtship and not endoreduplication, as a negative control. (Figure 5B and C) However, when I screened the eyes of the progeny, *GMR>GAL4* did not yield a suitable phenotype for the screening because the phenotype varied drastically from an orange eye color to the wild type eye color (red) to a dark red eye color. The change in eye color is concerning because each of the progeny possesses one allele of the *GAL4* driver and the one allele of the RNAi transgene with no other transgenes which may influence eye color. As a result, the eyes throughout all of the progeny ought to be the same eye yet after performing the experiment, this was clearly not the case. Variation of eye color makes judging more difficult

thus it would take longer to screen which was undesirable. In addition, there were variable ranges of 'rough' eye phenotype, the driver itself demonstrating a mild version of the rough eye phenotype without any *UAS>RNAi* which caused worry in regards to false positives for mitotic replication. *eyeless>GAL4* phenotypes did not vary as much but they were too extreme. The progeny either had little or no eyes or were pupal lethal and did not develop into a mature adult flies.

Learning that the tested eye *GAL4* drivers were not suited for my screening, I researched *GAL4* drivers expressed in wing discs, another mitotically cycling tissue. I decided to pre-screen *ms1096>GAL4* and *en>GAL4* because they had been utilized in other laboratories and were readily available in the lab. However, after crossing the *GAL4* drivers to the same RNAi transgene used in the previous eye screen, I did not observe any clear phenotype using either *GAL4* driver. There was no difference between a wild-type fly wing and the wings of the *GAL4/UAS* progeny.

At this point, I decided to ask for suggestions from Dr. Amanda Simcox because her research is involved in wing disc expression. She recommended three *GAL4* drivers to screen for mitotically replicating tissue and/or endoreduplicating tissue which is dpp^{blk1}*>GAL4, ptc>GAL4* and *71B*>GAL4. In my pre-screen, I compared the expression of these three *GAL4* drivers in endoreduplicating larval salivary gland tissue. To easily visualize the salivary glands, I genetically modified each of the three *GAL4* driver alleles to recombine with *UAS>GFPnls* alleles, which encodes Green Fluorescent Protein with a nuclear localization signal. The GFPnls is also downstream of an *UAS* sequence so it is activated by any Gal4 protein which binds to it.

I also added and tested *UAS>Dcr-2* (Dcr-2 under control of *GAL4/UAS)* to see if there was a significant difference in RNAi effects with and without Dcr-2. (10) Each driver was crossed to the same *UAS* lines used in the eye screen described above. Based on distributions of

UAS>GFPnls signal, all three drivers yielded similar phenotypes in larval salivary glands with and without *UAS>Dcr-2*. The experiment demonstrated that crosses with *UAS>Dcr-2* seemed to enhance the phenotypes. (Figure 4A and C and 5A) As a result, we decided to use *ptc>GAL4* recombined with *UAS>GFPnls* and *UAS>Dcr-2* based on the data from the pre-screen as well as previous driver characterization from past experiments. (4)

While screening the three salivary gland drivers, I learned from Dr. Simcox that dpp^{blk1}*>GAL4* was expressed in both mitotically replicating wing discs and endoreduplicating larval salivary glands. In consequence, I tested this driver against the same *UAS* lines from the previous pre-screening experiments using *GAL4* drivers expressed in the eye. While there were four pupal lethal progeny, the expressivity of the *GAL4* driver phenotype was more easily observed and it was very simple to score the varying strengths of the phenotype. (Figure 4B and D) Because the phenotype given by dpp^{blk1}*>GAL4* in larval salivary glands was comparable to the phenotype of *ptc>GAL4*, I decided to use dpp^{blk1}*>GAL4* also as the salivary gland driver. This driver seemed to perfectly suit the screen because I would be able to use one driver to screen two tissues which would expedite the screening process more by not having to cross *UAS>RNAi* lines to another driver and it would also save resources.

The Notch pathway has been shown to regulate the switch from mitosis to the endocycle in ovarian follicle cells. (11) After learning this, I realized that there is the possibility that some genes might only regulate the switch from mitotic replication to endoreduplication rather than regulating endoreduplication per se. For this purpose, I sought out a driver which would be expressed in ovarian follicle cells. During oogenesis, ovarian follicle cells switch from mitotic replication to endoreduplication from Stage 6 to 7 and are very well studied making them a high-priority tissue to screen. (Figure 5D) (12) My advisor, Dr. Maki Asano, and a former P.D. in our lab, Dr. So Young Park, had screened the expression of several ovarian follicle cell *GAL4* drivers

using *UAS>GFPnls*. (Figure 5E) Given this data, I decided to pre-screen *c355>GAL4* because expression of this driver begins from Stage 6 and continues onwards through oogenesis allowing me to screen for factors involved in the switch from mitotic replication to endoreduplication. I pre-screened *c355>GAL4* by crossing it to the *UAS>RNAi* lines used above. To screen for phenotypes, I dissected out the ovaries and visualized the nuclei with DAPI. Three crosses were pupal lethal but the other three gave a sterility phenotype as well as decreased size of follicle cells around the egg chamber during oogenesis. I did not combine this driver with *UAS>GFPnls*. It seemed too complicated to transgenically modify the driver since the gene is on the first chromosome. I did not carry out the pre-screen with *UAS>Dcr-2* either because, given the amount of progeny which were lethal, there was a high chance that co-expressing Dcr-2 would not yield adults or adults without ovaries.

At this point in the pre-screening, I realized both *GAL4* drivers I had selected thus far were expressed in epithelial cells. dpp^{blk1}*>GAL4* is expressed in larval salivary glands and *c355>GAL4* in ovarian follicle cells. As I am primarily interested in 'generic' endoreduplication factors that are required in the various cell types, I decided to screen germline nurse cells. I was already using ovaries to screen follicle cells and it would give me another assay in which to screen for endoreduplication-specific factors. The driver I chose to use was *VP16nos>GAL4* because it is a well characterized and has been successfully tested. (Figure 5F) (13) Using *VP16nos>GAL4*, I compared the effects of *UAS>Dcr-2* in the nurse cells. We dissected out and visualized the ovaries of the progeny with *UAS>GFPnls* as well as stained them with DAPI. The results between the two varied a lot. The transgenic *VP16nos>GAL4* line without *UAS>Dcr-2* did not yield any phenotype while the transgenic line containing *UAS>Dcr-2* did. For instance, some yielded ovaries with increased numbers of nurse cells and with smaller nuclei in Dcr-$2^{(+)}$ crosses while Dcr-$2^{(-)}$ did now. This demonstrates inhibition of endoreduplication.

Genome-wide Screening of Drosophila

Using three drivers identified from the pre-screening, I moved on to the 'pilot' screen. The first, used to screen larval salivary glands and adult wing tissue, is *dpp^{blk1}>GAL4* which also encodes *UAS>Dcr*-2 and *UAS>GFPnls*. The second is *c355>GAL4* and it is used to screen ovarian follicle cells. The third is *VP16nos>GAL4* transgenically modified to carry *UAS>GFPnls* and *UAS>Dcr-2*. I grew five sublines for *dpp$^{blk1>}$GAL4* and *c355>GAL4* and ten sublines for *VP16nos>GAL4* because *VP16nos>GAL4* viability is not as high as the other two.

Drosophila development usually occurs in the span of 10 days. (Figure 6A) To increase efficiency, I developed a screening schema. (Figure 6B-E) I order 160 *UAS>RNAi* lines from Vienna Drosophila RNAi Center every week. They are in sequential order according to their line number. When the lines arrived, I transfer them into new vials. Before the *UAS>RNAi* lines arrive, my peers and I collect virgins from the *GAL4* driver sublines. Next, we aliquot and cross *GAL4* driver virgins to males containing the *UAS>RNAi* transgene. After three to four days, I discard the parents of each cross to prevent confusion between parents and progeny when the progeny fully emerge. After five to six days, I screen larval salivary glands of the *dpp^{blk1}>GAL4* crosses. After 10 or 11 days, I screen the wing tissue of the progeny of the *dpp^{blk1}>GAL4*. On the same day, I also aliquot three to four male and three to four female progeny from the cross *c355>GAL4* and *VP16nos>GAL4* crosses into new vials. I check for fertility after three to four days.

Scoring phenotypes was done by using a system refined from the range of phenotypes in pre-screening data. The *UAS>RNAi* lines which we find to be the most interesting for *dpp^{blk1}>GAL4* are those which yielded a salivary gland medium strength or stronger phenotype (i.e. salivary glands which were at least less than half of the control salivary gland size) was with

no, or a very mild, wing phenotype. The lines which were most interesting for *c355>GAL4* and *VP16nos>GAL4* were those which yielded progeny that demonstrated semi-sterile (severe reduction in fertility) or sterile phenotype.

After scoring the ~1280 RNAi lines, we decided to further characterize the positive candidates received from the pilot screen. To do so, we utilized Fluorescence Activated Cytometry Sorting (FACS) for follicle and nurse cell candidates and clonal analysis for the salivary gland candidates. (Figure 11 and 12) The follicle and nurse cell positive candidates were screened using the fertility. The FACS analysis allows me to look at the ploidy of the follicle and nurse cells and I have a clearer idea if the sterile phenotype is caused by a shift in ploidy due downregulation of endoreduplication-specific factors or those involved in the switch from mitosis to the endocycle, or if it is attributed to another process not related to endoreduplication. The clonal analysis allows me to distinguish if change in cell nuclei size is occurring in a cell autonomous manner. Additionally, it allows me to separate candidates specifically involved in endoreduplication from those which may be involved in other cellular processes affecting cell size such as metabolism.

I ran a FACS analysis on the 58 candidates identified though the follicle cell screen and found that 20 of them yielded a shift in ploidy. This was determined by looking at trends in the data and establishing a cutoff point. I saw that there were several crosses which revealed a slight shift in ploidy, a 10% difference in the 16N peak from the control. However, these did not show a significant shift otherwise so I decided to separate the more interesting candidates, those which showed a decrease in the 16N peak by at least 10% complemented by an increase in the lower ploidy peaks, from those which do not yield a significant enough phenotype. While the FACS analysis for the candidates identified from the follicle cell screen went smoothly, the nurse cell FACS analysis required a lot of troubleshooting. Since I was not receiving constant profiles for

the nurse cell candidates, I referenced the original protocol and adjusted conditions accordingly. The primary experimental conditions that I changed were the centrifuge time and speed as well as the concentration of collagenase. *Calvi BR and Lilly MA* use 14000 rpm in their protocol for 20 minutes. In addition, they use 5 mg/ml collagenase. I changed one variable in each conditioning experiment. When I changed the centrifuge time to 20 minutes, this did not affect the data at all. Next, I adjusted the centrifuge speed from 25000 rpm to 14000 rpm but this produced even worse results. However, when I changed the concentration of the collagenase from 10 mg/ml to 5 mg/ml, I saw much better results. In consequence, I decided to use 5 mg/ml collagenase, centrifuging at 25000 rpm for 20 minutes. I chose 20 minutes since it is more true to the original protocol and did not seem to matter if the samples were centrifuged for 20 minutes or an hour.

Discussion

Out of the 1280 genes screened in the pilot screening, 216 genes were identified as having possible function in endoreduplication and/or the switch from mitosis to endoreduplication. (Figure 7) Three of these are of particular interest because they all demonstrated possible involvement in endoreduplication in all three tissues. *Nucleoporin 54 (Nup54)* is involved in the creation of cytoplasmic compartments in the nuclear envelope. *Whgsc2* is a gene which has not been characterized very much. *Cp7Fa* is an uncharacterized gene which does not have any known mouse of human homologues. Out of all of these, I have only characterized *Nup54* via the FACS and have yet to do a clonal analysis on it. The *UAS>RNAi* lines for *Whgsc2* and *Cp7Fa* did not reproduce well and were lost. We plan to reorder them and characterize them as soon as possible. Looking at DNA/Endo reports for

Nup54 has revealed that *Nup54* expression is high in early *Drosophila* embryonic development, moderately expressed throughout larval and pupal stages, and then is highly expressed again in adult males and females. Moreover, it is moderately expressed in larval salivary glands and in ovaries. Moreover, it is involved with protein importation, more specifically with direct interaction with the NLS-bearing substrate, into the nucleus. There are also known orthologs in *Caernorhabditis elegans*, *Xenopus tropicalis*, and *Mus musculus*. When taking the screening data into account and after characterizing using bioinformatics, this gene has high potential to be involved in endoreduplication. There is one available allele from Bloomington Stock Center and this will be analyzed using more hypothesis-driven methods.

Many candidates demonstrated a severe salivary gland phenotype and did not develop into viable adults. In addition, many *UAS>RNAi* lines when crossed to the follicle cell *GAL4* driver, *c355>GAL4*, did not yield viable adults and prevented me from performing a fertility check or FACS analysis. This lethality may be attributed to a lack of specific gene expression. For example, *His2B/CG33868* is a gene needed for histone integrity. If it is compromised, the DNA in cells will be mutated or destroyed causing the specific developmental disc to not develop properly. If this occurs, the organism will not be able to become a mature, adult fly. Another possibility for lethality might be due to metabolism genes. Many cells, especially the enormous larval salivary gland cells, require large amounts of nourishment. If an RNAi line targets expression of a gene, like *CG9512/Glucose Dehydrogenase*, essential towards glucose metabolism, cells will not be able to develop properly. This might cause cells, such as salivary gland cells which are normally very large, to seem smaller not because of a defect in endoreduplication, but due to a metabolism defect.

To find what types of functions the candidates from the primary screening are assigned, I

ran a PANTHER (Protein ANalysis THrough Evolutionary Relationships). (Figure 8, 9 and 10) When interpreting the data from PANTHER, I have to keep in mind several aspects which might change the way I interpret the PANTHER data. One is that the candidates identified from the salivary gland and nurse cell screen have a high number of RNA binding proteins, which are likely to be involved in translation. This interpretation is consistent with the fact that salivary gland cells and nurse cells are much larger than follicle cell so that they require a higher metabolism. These proteins are probably not involved in an endoreduplication-specific manner and will be excluded through the secondary screening. A particular intriguing category of proteins are the DNA-binding proteins which yielded comparable numbers of candidates from each tissue screen. These proteins are interesting because DNA-binding proteins may include transcription factors which may be involved in endoreduplication-specific regulation. However, since the number of candidates included in the PANTHER is small, it is hard to draw definitive conclusions. As the screening continues, the data will become more reliable because there will be a larger pool of candidates.

There are three goals for the future. The first is to complete the secondary screening for the candidates identified from the salivary gland and nurse cell screens. There are still 90 nurse cell candidates which have not been screened by FACS Analysis. I expect that the percentage of remaining number of candidates will be similar to the candidates identified in the follicle cell screening. Out of 58 potential candidates identified from the follicle cell fertility screen, 20 remain demonstrating that only 40 percent of the lines remain. However, 10 lines could not be screened because the flies died. As a result, this number might increase to about 50 percent. If this holds true for the nurse cell candidate FACS analyses, then out of 103, ~40 would remain. These numbers of candidates are much more manageable than the amount of candidates identified from the fertility screens. Conditioning and execution of the clonal analysis for the

candidates from the larval salivary gland screen would be concurrent with the nurse cell candidate FACS. Having never done the clonal analysis myself, I cannot approximate how many candidates I expect will remain but if it is a similar percentage to the candidates identified from the follicle cell fertility screen, then I would hope that ~40 out of 109 salivary gland candidates would remain.

The second goal is to continue screening the remaining ~10000 *UAS>RNAi* lines using the same schema as the pilot-screen (Figure 6). Since there is a set protocol and not much conditioning needs to be done, it will be performed in a more efficient manner. Out of the 1280 lines screened in the pilot screening, 216 were identified as possibly being relevant to endoreduplication in some way. Given the percentage above, I would expect that after screening the remaining *UAS>RNAi* lines, ~1600 would be identified as possibly being involved in endoreduplication.

The final goal is to further characterize the remaining genes from the secondary screen. For instance, I would begin hypothesis-driven experiments with *Nup54* to further characterize its specific interactions with endoreduplication in *Drosophila*. Another possibility would be to study the identified genes in other organisms which possess greater homology with humans such as mice. If I use mice, I would be able to use the Cre-LoxP system which allows tissue-specific knockouts. With this, I could knockout the candidate gene in endoreduplicating mice cells such as giant trophoblast cells or hepatocytes. If the gene is necessary for growth, then cells which would normally endoreduplicate would theoretically be smaller in size and contain lower DNA content than their wild-type counterparts.

Appendix

As I reflect on my undergraduate experience in the field of research, I realized that I had a very naïve sense for what was expected. I had two major influences: Human Genome Project and the Ebola virus. I was in awe of the daunting task the scientists faced when attempting to sequence an entire genome. As a result, I felt challenged to overcome obstacles even greater than the Human Genome Project. Even more so, I felt a huge urge to help people and desired to find treatments for sicknesses ranging from the common cold to terrible epidemics like the Ebola virus. When I first began research, I understood that human beings have gained much insight from doing research and that these advances in science have been very beneficial to many people. However, I was not ready for what research is and what was involved. If I was to summarize my overall experience in research now, I would say it is one of the time-consuming but enlightening and rewarding activities in which I have ever participated.

Reflecting on my experience, I believe the first big shock for me was the realization that there are actually two different types of research: hypothesis-driven research and large scale input research. When I imagined research, I was under the impression that all of it was hypothesis-driven. Research was the process of discovering how to treat diseases or the journey of discovering something new and innovative which would immediately resound throughout the world. The thought of looking for new interactions and genes through large scale input research never occurred to me. As a result, when introduced to the concept of a genome-wide screening, I was puzzled. The screening falls into the large scale input-type project and serves to identify factors specifically involved in endoreduplication and/or the switch form mitosis to endoreduplication. After using the schema (Figured 5) to identify possible candidates and using the secondary screening to further characterize them (Figure 11 and 12), the remaining candidates will be tested through more hypothesis-driven research. This ties into future goals and

will be discussed later.

Another lesson that I quickly learned was that requires immense amounts of work. I slowly began to realize that as a researcher, I needed to have a very different timetable from other professions which fall into the typical workday. I found that the model organism, in my case *Drosophila,* dictate the researcher's timetable, not the other way around. However, when I first began, I did not understand this concept. There were many times when I would set up an experiment only to fail half-way through because I did not time my experiments correctly. One persistent example would be when I had to collect female virgins (females which have yet to mate with a male). The developmental lifecycle of *Drosophila* occurs in the span of 10 days. When females emerge, there is a short period of time which I can collect them and be sure they are virgins. If the flies emerged during the weekend, I would wait until the weekday to collect the females, hoping that there would still be virgins to collect. As a result, I would not be able to gather many virgins and if I did not have enough, I would have to wait another 10 days until the next generation to collect more. Even though this seems to be inconsequential, the misconception of time I possessed while collecting virgins could pose terrible consequences in reference to the screening. One severe example would be screening progeny which do not have the correct genotype causing collection of incorrect data. These experiences delayed my research a little but allowed me to gain a greater understanding of how to budget my time. Moreover, it solidified the concept that I was not at the center of my world and that sacrifice is sometimes needed for advancement.

Aside from learning the methodology and how to budget my time, research has taught me how to present my research and how to write as well. For instance, I presented in casual settings such as general lab meetings and on multiple occasions would present my research to other

Drosophila labs. Additionally, I presented at research forums which allowed me create a poster and refine how I would present my project in such a setting. Thanks to an immense amount of guidance and my hard work, I even received a 2nd award for the 2011 Denman Undergraduate Research Forum. The final aspect of research other than experiments which I have learned is how to write in a scientific setting. Having applied for many scholarships and fellowships during my undergraduate career, I have been graced with one scholarship and am also a recipient of the Pelotonia Undergraduate Research Fellowship in 2010 and 2011. The financial support from these awards has allowed me to focus more time on my research because I did not have to worry about living expenses. Moreover, I was asked to participate in the 2011 Pelotonia Symposium. Through this, I acquired the knowledge of orally conveying my research in a more formal setting. For the symposium, I produced a presentation where I presented my research to many distinguished scientists in the field of cancer research. In addition, as a Pelotonia fellow, I was prompted to participate in certain events which gave me perspectives on research I was never aware of such as the Pelotonia, the bike ride to end cancer. After hearing many tragic cancer patient stories, I realized the importance my research could potentially have on the world. This spurred me forward and gave me even greater enthusiasm towards my research.

Acknowledgments

I would like to thank my advisor, Dr. Maki Asano, for all the guidance she has given. This would not have been possible without her continual trust and dedication towards my education. I would also like to thank Dr. Amanda Simcox for her guidance and discernment when I was faced with seemingly insurmountable obstacles. Moreover, Dr. Gustavo Leone has also been a dedicated mentor and has aided me in many aspects regarding the screening. Next, I would like to thank Dr. Lindsey Kent, Dr. Jaehong Park, Dr. So Young Park, and Dr. Takayuki Uchida for their guidance and their investment in my education. In addition, I would like to thank Dr. Tamar Rudavsky for her flexibility and her grace when asked to be on my honors thesis oral defense committee. Also, I would like to thank the Flow Cytometry Lab personnel, namely Daniel Lundy, Bryan McElwain, and Katrina Moore located in the Biomedical Research Tower for their cooperation and flexibility in regards to the FACS Analysis.

Finally, I would like to thank Alisar Aboudiab, Gregory Adams, Greeshma Allareddy, Kelsey Durbin, Jeffey Frank, Mike Freeze, Amrita Jagpal, Aiko Kondo, Kyle McClintock, Aisha Nurr, Julianna Paterra, Sarah Paul, Edwin Rice, Kaitlin Rubinic, Fatima Siddiqui, Emily Spetka, Jackie Stuber, Nurefsan Talayhan and Jay Vance and for their hard work and dedication and patience.

Figures

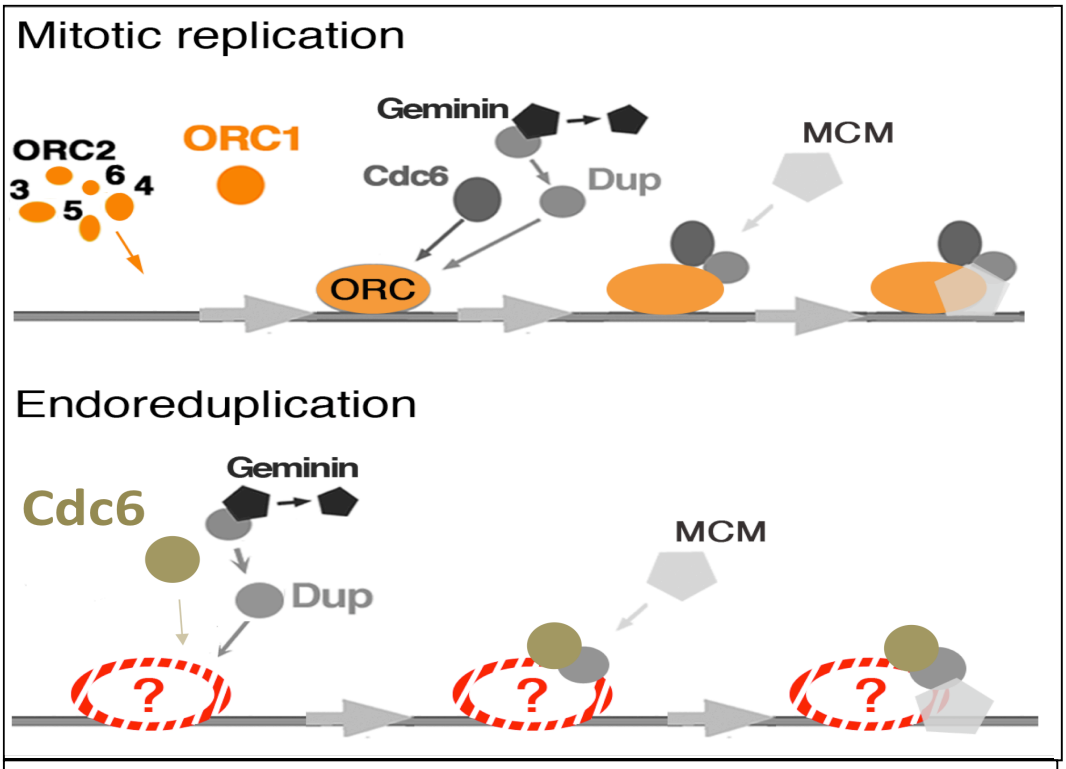

Figure 1 – Formation of the Pre-Replication Complex in Mitotic Replication and Endredouplication

In mitotic replication, the Origin Recognition Complex (ORC) is necessary to form the pre-Replicative Complex (pre-RC) which initiates DNA replication. However, in endoreduplication, ORC is dispensable. Another protein or set of proteins is responsible for the formation of the pre-RC.

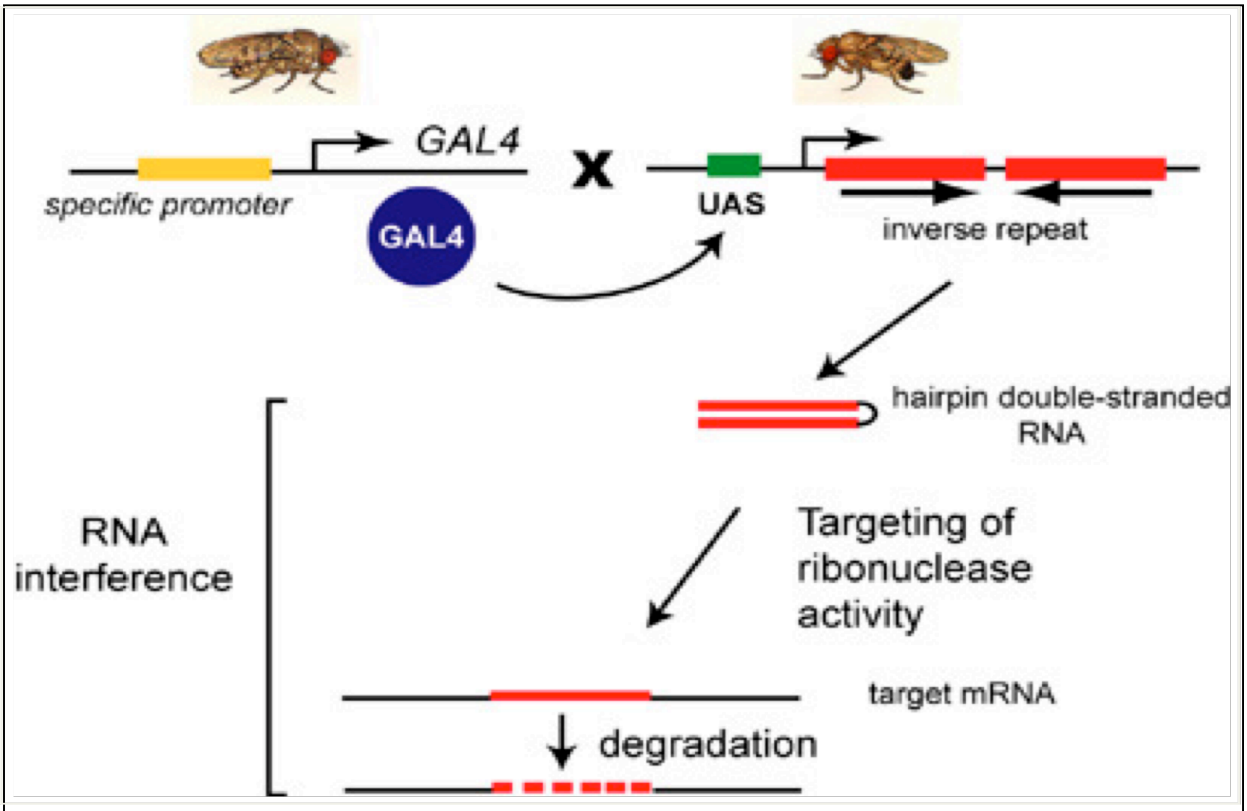

Figure 2 – The GAl4/UAS system with RNAi

The GAL4 protein is transcribed under the control of a tissue-specific promoter. It binds to the Upstream Activating Sequence (UAS) and causes transcription of mRNA. The mRNA transcribed is an inverse repeat which will form dsRNA and be integrated into the RISC complex. The RISC targets homologous mRNA for degradation.

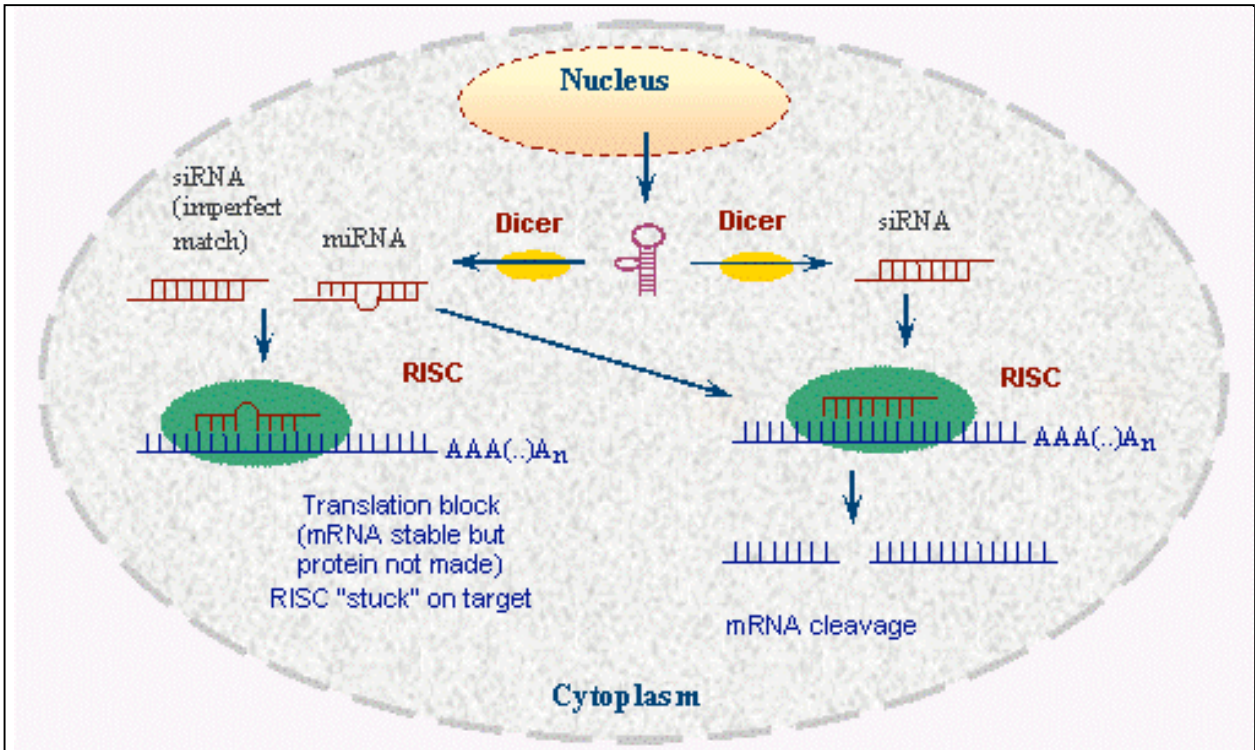

Figure 3 – RNAi Pathway

Dicer targets and cuts both siRNA and miRNA. The cut siRNA and miRNA are incorporated into the RISC complex which targets homologous sequences of mRNA. If miRNA, is incorporated into the RISC, translation of mRNA is inhibited. If siRNA is incorporated, mRNA is cleaved. Both methods serve to downregulate protein expression,

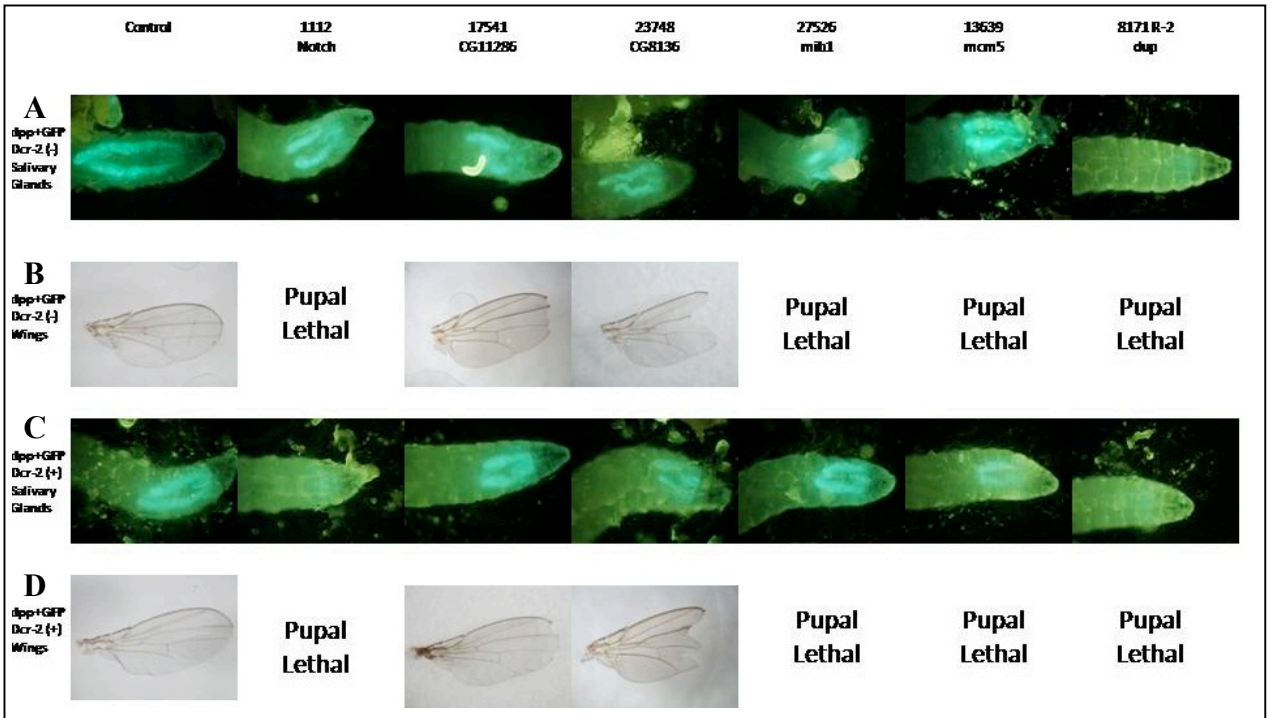

Figure 4 - Pre-screening of larval salivary glands and wings using *dpp^{blk1}>GAL4* with and without *UAS>Dcr-2*

Figures A and B. *UAS>Dcr*-2$^{(-)}$ crosses to *dpp^{blk1}>GAL4+UAS>GFPnls.*
Figures C and D. *UAS>Dcr*-2$^{(+)}$ crosses to *dpp^{blk1}>GAL4+UAS>GFPnls*.
The phenotypes from C are much more severe than the phenotypes from A. The phenotypes from D are also more severe than those from B. This demonstrates that *UAS>Dcr-2* enhances RNAi expression.

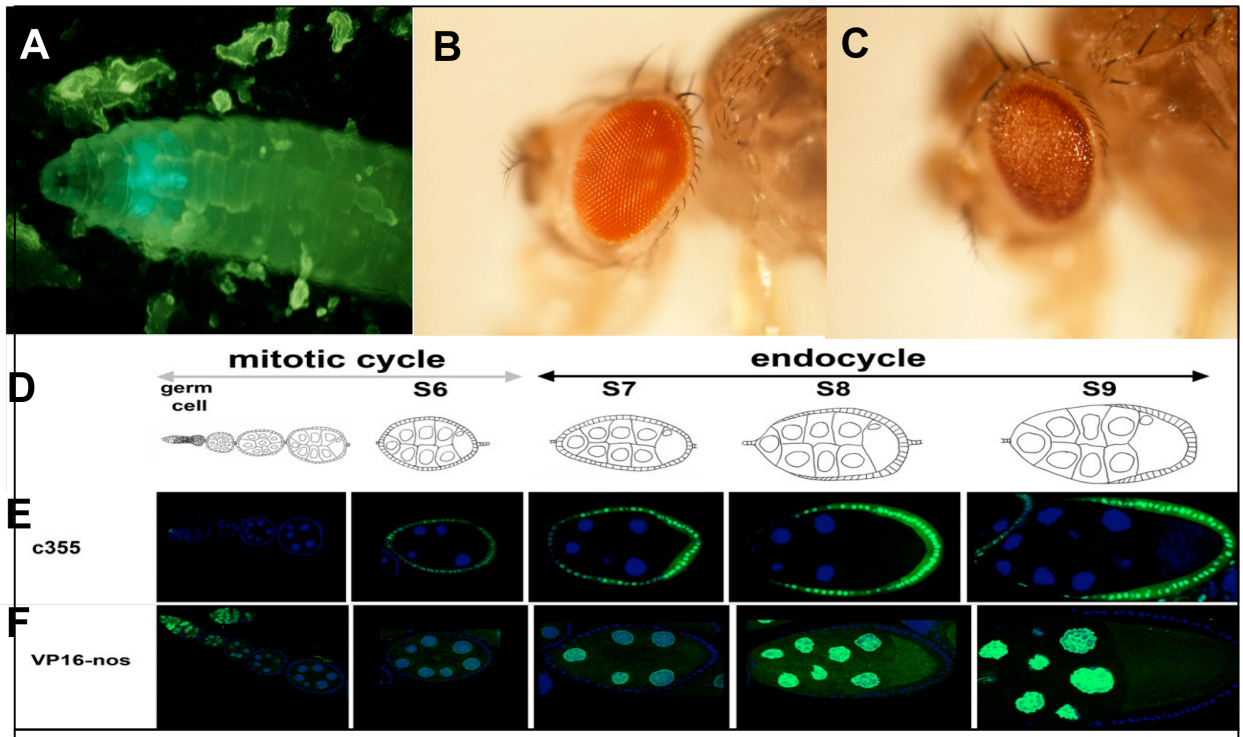

Figure 5 – Expression of *dpp*^{blk1}*>GAL4, GMR>GAL4, c355>GAL4,* and *VP16nos>GAL4* and *Drosophila* oogenesis.

Figure A. *dpp*^{blk1}*>GAL4* salivary gland expression utilizing *USA>RNAi* in conjunction with *UAS>Dcr-2* and marked with *UAS>GFP*.
Figure B. Wild-type eye phenotype.
Figure C. *GMR>GAL4* expression in the eye (rough eye phenotype).
Figure D. A cartoon depicting the stages of oogenesis through stage 9 in addition to the switch from mitotic replication to endoreduplication.
Figure E. Expression of *c355>GAL4*, the follicle cell driver, marked with GFP.
Figure F. Expression of *VP16nos>GAL4,* the nurse cell driver, marked with GFP.

Figures D-F – So Young Park

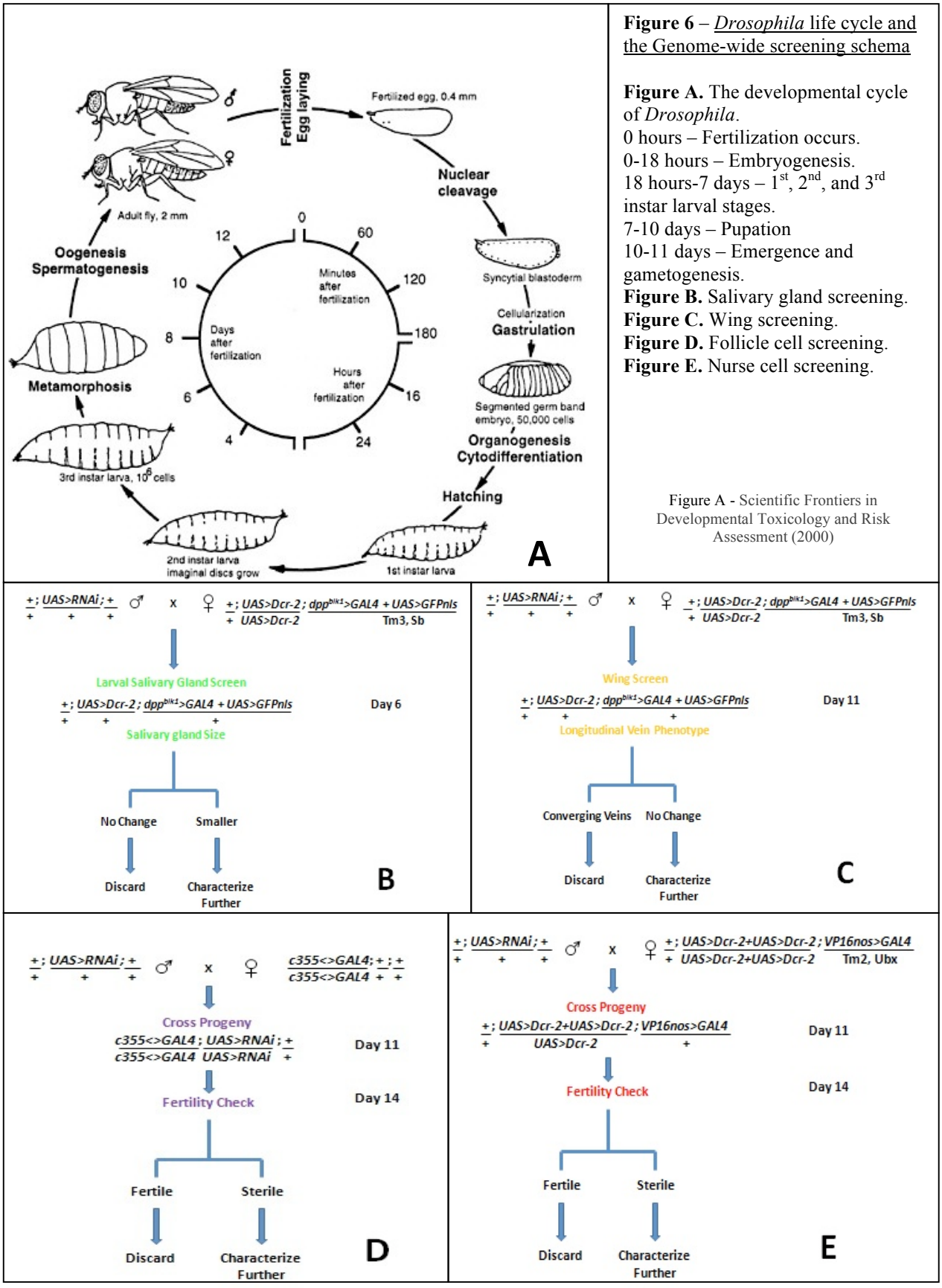

Figure 6 – *Drosophila* life cycle and the Genome-wide screening schema

Figure A. The developmental cycle of *Drosophila*.
0 hours – Fertilization occurs.
0-18 hours – Embryogenesis.
18 hours-7 days – 1st, 2nd, and 3rd instar larval stages.
7-10 days – Pupation
10-11 days – Emergence and gametogenesis.
Figure B. Salivary gland screening.
Figure C. Wing screening.
Figure D. Follicle cell screening.
Figure E. Nurse cell screening.

Figure A - Scientific Frontiers in Developmental Toxicology and Risk Assessment (2000)

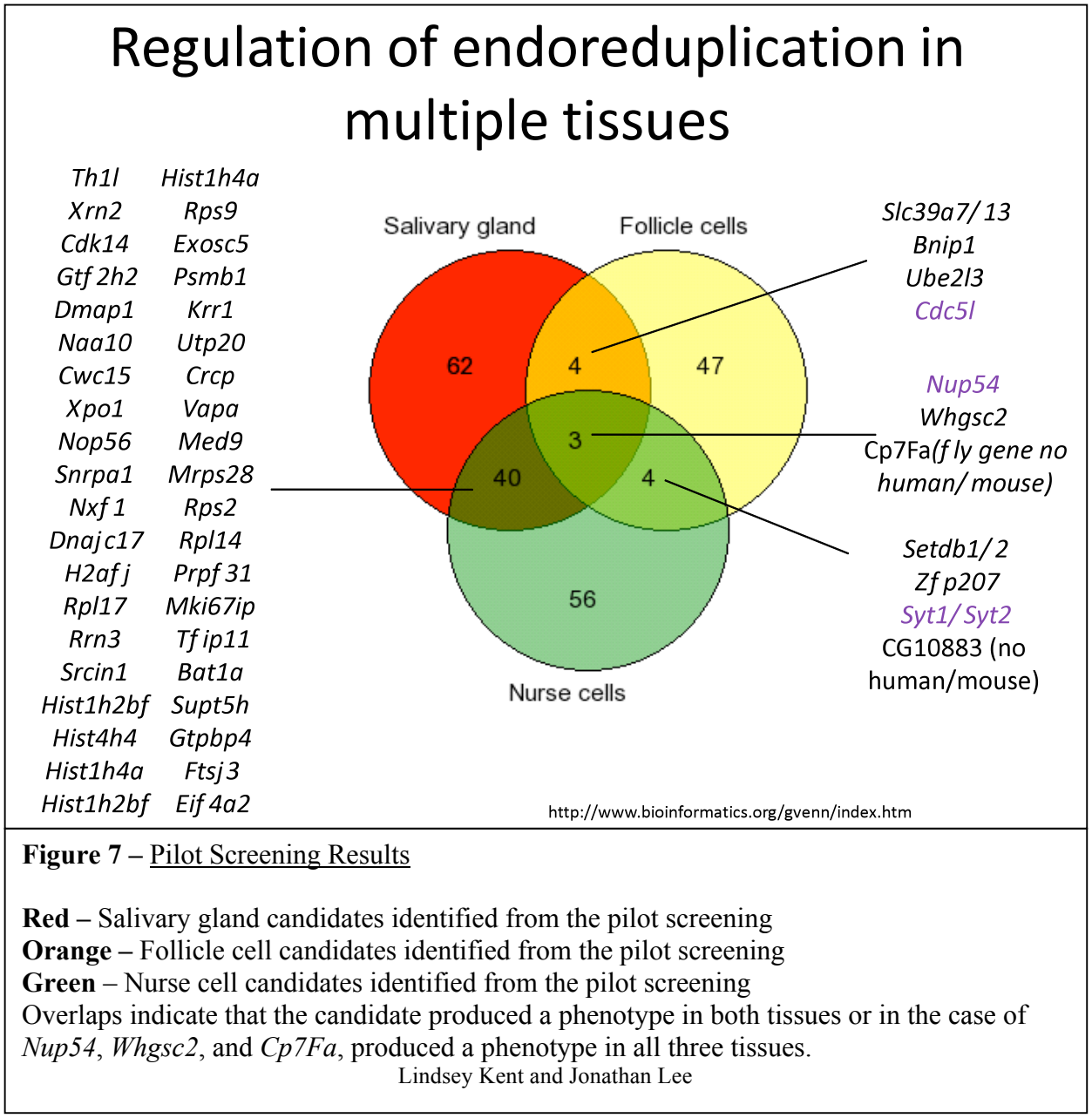

Figure 7 – Pilot Screening Results

Red – Salivary gland candidates identified from the pilot screening
Orange – Follicle cell candidates identified from the pilot screening
Green – Nurse cell candidates identified from the pilot screening
Overlaps indicate that the candidate produced a phenotype in both tissues or in the case of *Nup54*, *Whgsc2*, and *Cp7Fa*, produced a phenotype in all three tissues.

Lindsey Kent and Jonathan Lee

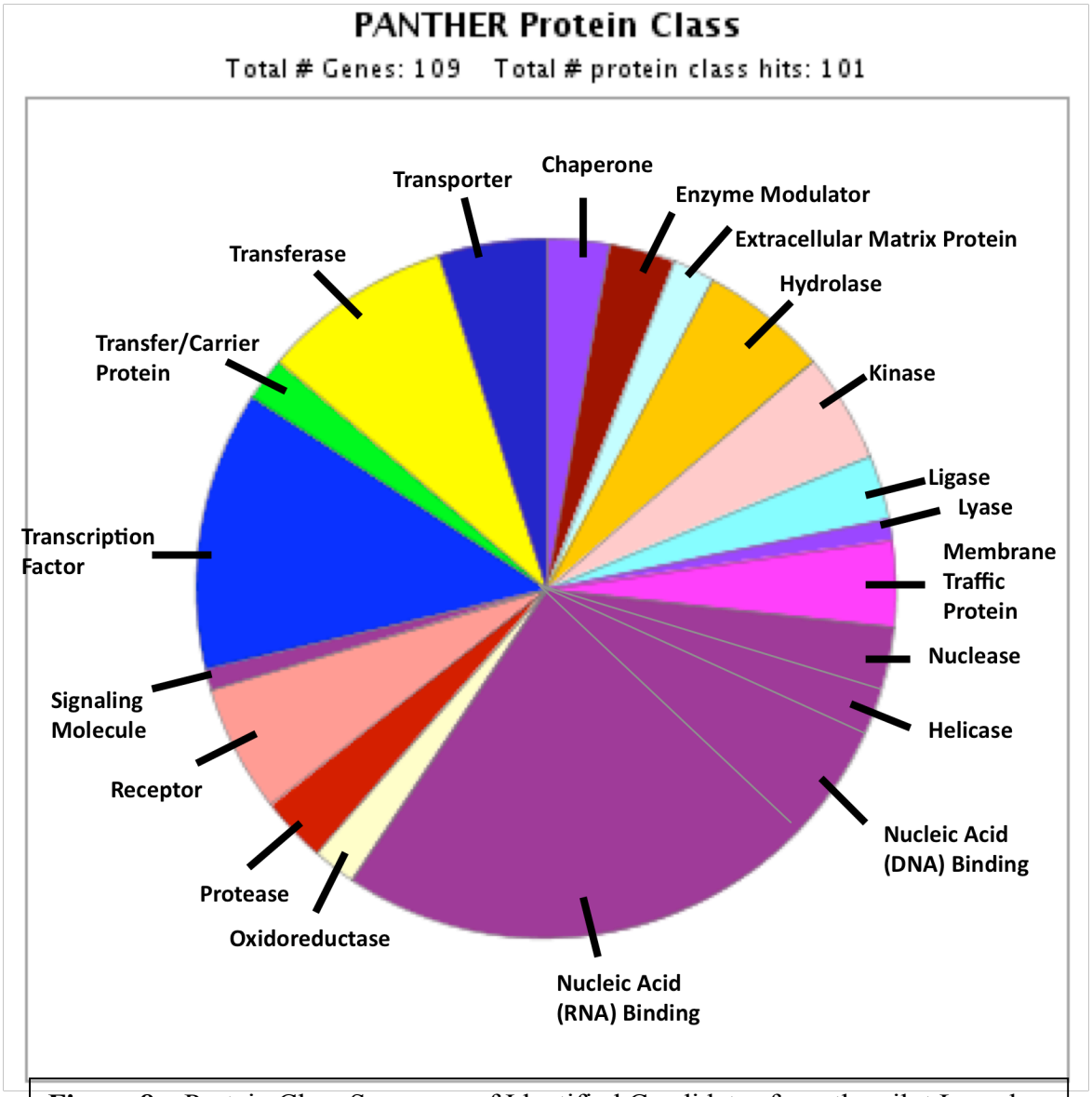

Figure 8 – Protein Class Summary of Identified Candidates from the pilot Larval Salivary Gland Screen

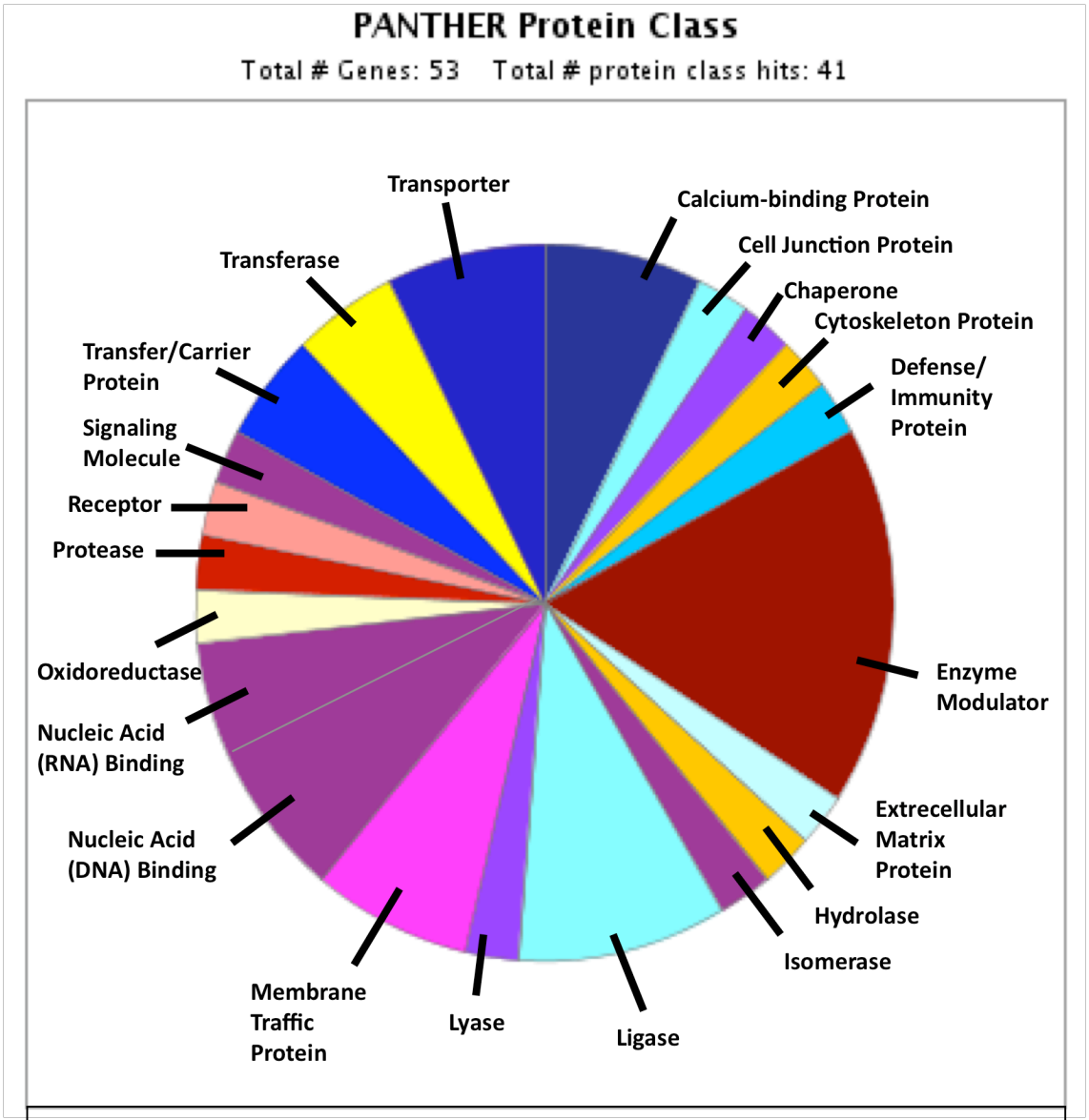

Figure 9 – Protein Class Summary of Identified Candidates from the pilot ovarian follicle cell screen

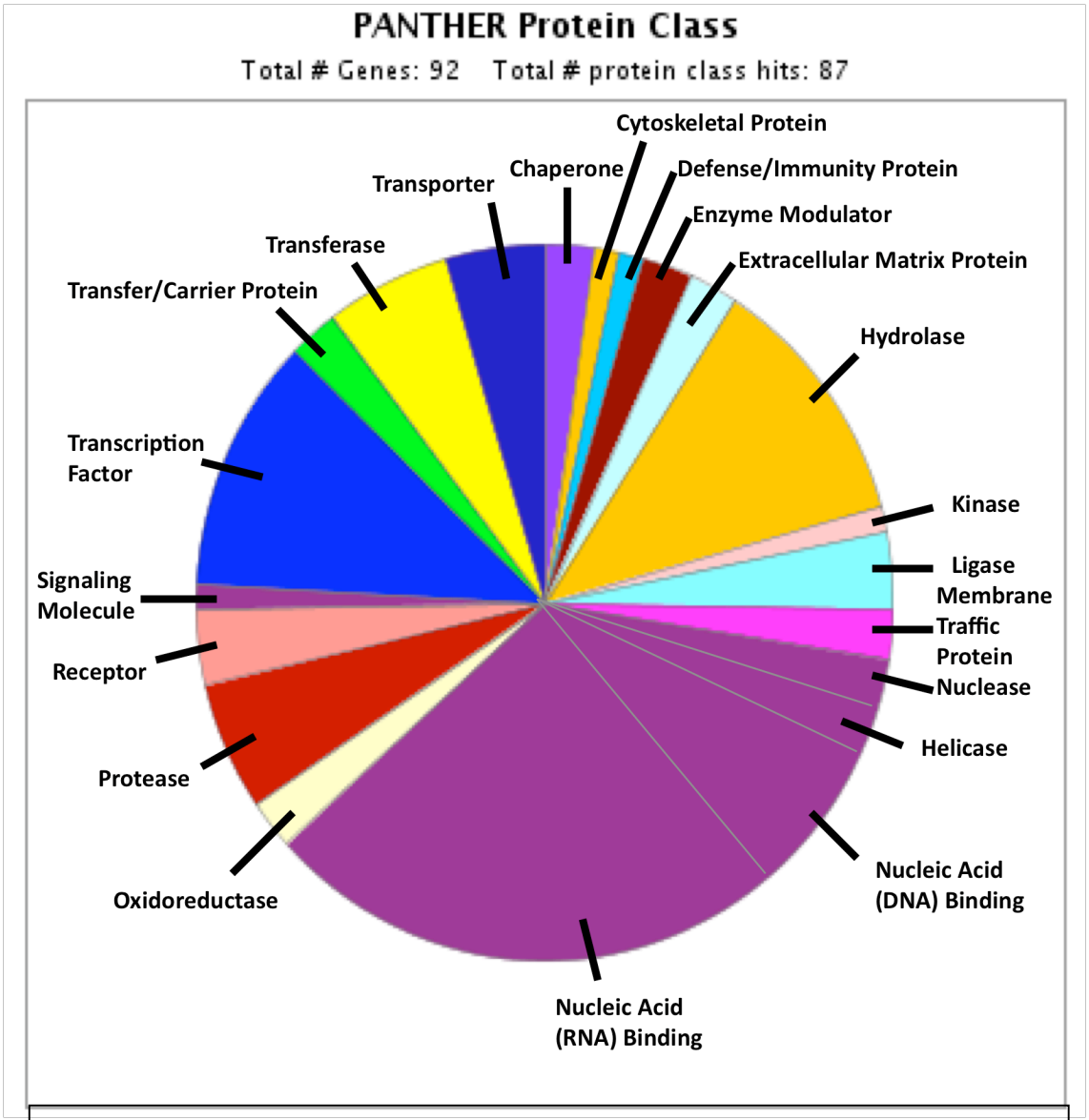

Figure 10 – Protein Class Summary of Identified Candidates from the pilot ovarian nurse cell screen

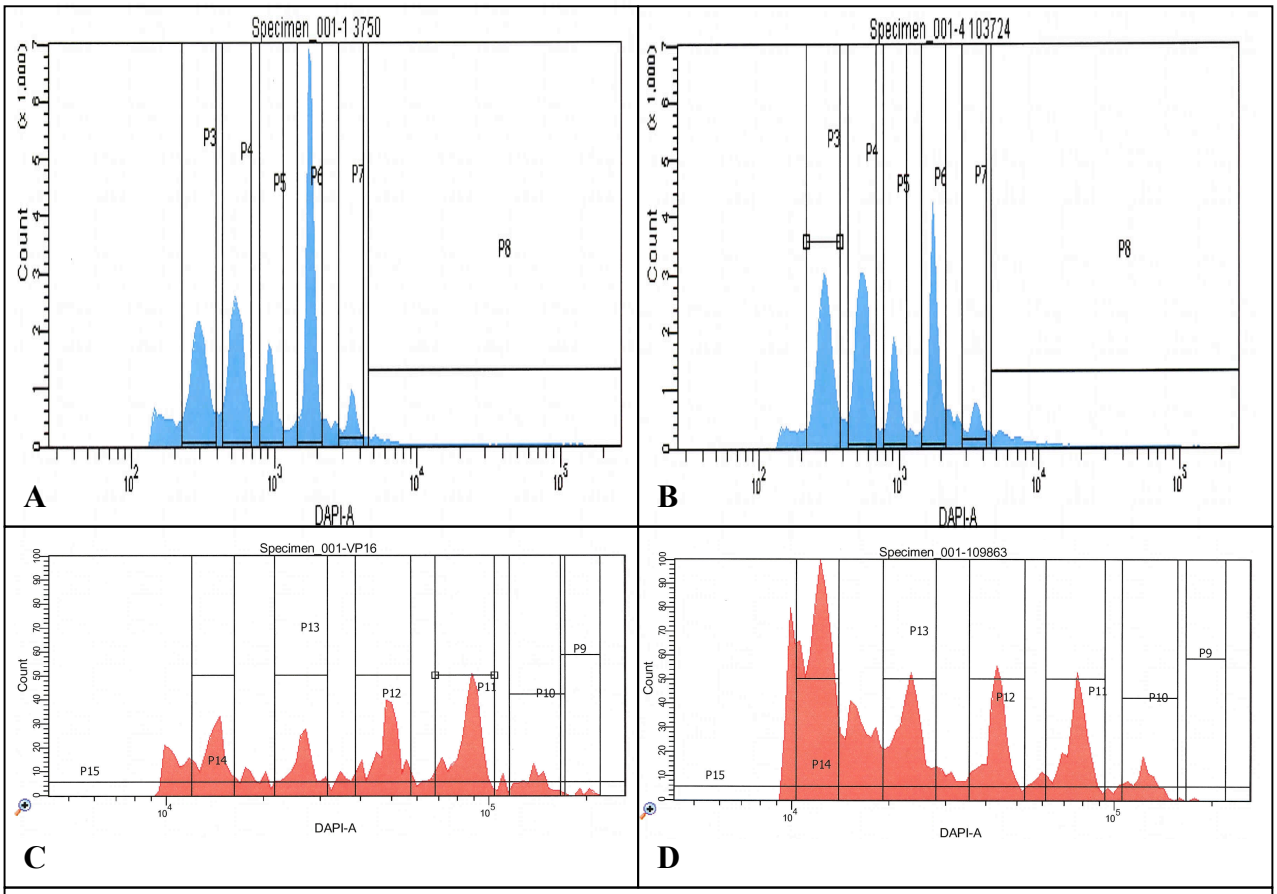

Figure 11 – *c355>GAL4* and *VP16nos>GAL4* FACS Analysis Profiles

Figure A and B. FACS analysis for *c355>GAL4* which is expressed in ovarian follicle cells. A is the *GAL4* driver itself and B is a *Nup54*. There is a clear shift downwards in ploidy and the 16N peak is much smaller than the control.
Figure C and D. FACS analysis for *VP16nos>GAL4* which is expressed in germline nurse cells. C is the *GAL4* driver itself and D is *hikaru genki*. The amount of cells counted is different but there is still a clear shift downwards in ploidy.

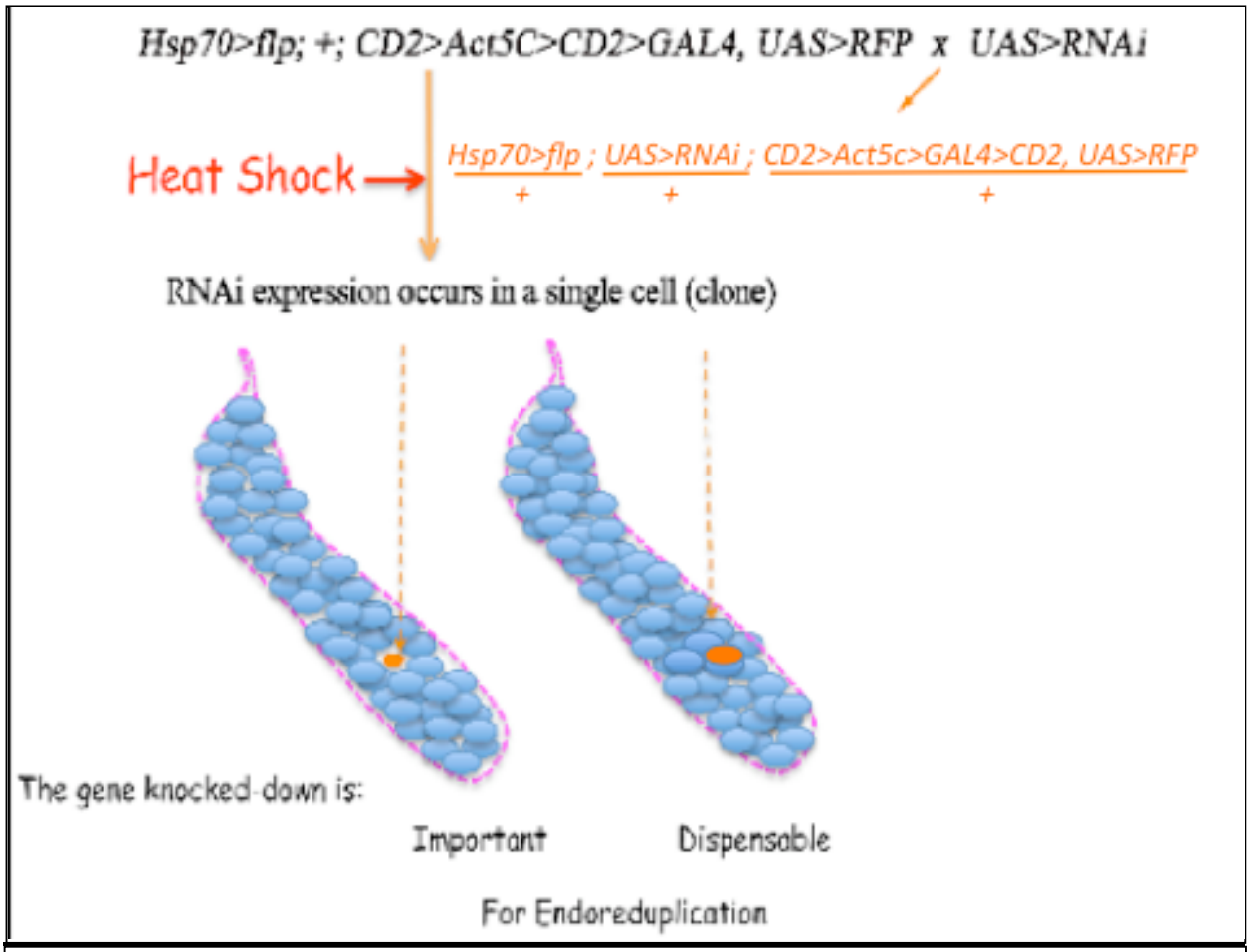

Figure 12 – Inducible clonal analysis of a gene using an RNAi transgenic line
Heat shock will induce flp recombinase expression which removes CD2 and thus activates Actin5C promoter to drive the GAL4. Clones will be marked by RFP.

Maki Asano

References

1. Edgar BA, Orr-Weaver TL (2001) Endoreplication cell cycles: More for less. *Cell* 105: 297–306.

2. Park SY, Asano M (2008) The origin recognition complex is dispensable for endoreplication in *Drosophila*. *PNAS* 105: 12343-12348.

3. Chauffert B et al. (2012) Tumor cells can escape DNA-damaging cisplatin through DNA endoreduplication and reversible polyploidy. *Cell Biol Int* 32: 1031-1043.

4. Brand AH, Perrimon N (1993) Targeted gene expression as a means of altering fates and generating dominant phenotypes. *Dev* 118: 401-415.

5. Dickson et al (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448: 151-157.

6. Galko et al (2010) A targeted *UAS-RNAi* screen in *Drosophila* larvae identifies wound closure genes regulating distinct cellular processes. *Genetics* 186: 943-957.

7. Mello C et al (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391 (6669): 806–811.

8. Kramer JM, Staveley BE (2003) GAL4 causes developmental defects and apoptosis when expressed in the developing eye of *Drosophila melanogaster. Genet Mol Res* 2:43-47.

9. Callaerts P et al (2009) Mutational analysis of the eyeless gene and phenotypica rescue reveal that an intact Eyeless protein is necessary for normal eye and brain development in *Drosophila*. 34: 503-512.

10. Taira K et al (2004) Overexpression of Dicer enhanced RNAi-mediated gene silencing by short-hairpin RNAs (shRNAs) in human cells. *Ox Jour* 48: 191-192.

11. Deng et al (2008) Regulation of the endocycle/gene amplification switch by Notch and ecdysone signaling. *J Cell Biol* 182:885-896.

12. Asano M (2009) Endoreplication: the advantage to initiating DNA replication without the ORC? *Proc Natl Acad Sci* 3: 173-175.

13. Kadonaga JT et al (1993) Potentiatoin of RNA polymerase II transcription by Gal4-VP16 during but not after DNA replication and chromatin assembly. *Genes Dev* 7:1779-1795.

14. Calvi BR, Lilly MA (2004) Fluorescent BrdU labeling and nuclear flow sorting of the Drosophila ovary. *Methods Mol Biol* 247: 203-213.